

## Visualisasi Kualitas Penyebaran Informasi Gempa Bumi di Indonesia Menggunakan *Twitter*

Mira Chandra Kirana<sup>1\*</sup>, Nanda Putra Perkasa<sup>2\*\*</sup>, Muhammad Zainuddin Lubis<sup>3\*\*\*</sup>, Maidel Fani<sup>4\*\*</sup>

\* Teknik Informatika, Politeknik Negeri Batam

\*\* Teknik Multimedia Jaringan, Politeknik Negeri Batam

\*\*\* Teknik Geomatika, Politeknik Negeri Batam

[mira@polibatam.ac.id](mailto:mira@polibatam.ac.id)<sup>1</sup>, [nanda.putra.perkasa@gmail.com](mailto:nanda.putra.perkasa@gmail.com)<sup>2</sup>, [zainuddinelubis@polibatam.ac.id](mailto:zainuddinelubis@polibatam.ac.id)<sup>3</sup>, [maidelfani@polibatam.ac.id](mailto:maidelfani@polibatam.ac.id)<sup>4</sup>

### Article Info

#### Article history:

Received 14-04-2019

Revised 08-05-2019

Accepted 10-05-2019

#### Keyword:

Gempa Bumi,  
Visualisasi,  
Penyebaran Informasi,  
Kualitas Informasi

### ABSTRACT

Indonesia merupakan salah satu negara dengan tingkat bencana alam yang cukup tinggi, seperti gempa bumi, sehingga penyebaran informasi mengenai peringatan dini sangat penting. Media sosial *Twitter* menjadi salah satu tempat penyebaran informasi mengenai peringatan dini bencana alam gempa bumi di Indonesia melalui akun BMKG. Namun belum diketahui kualitas penyebaran informasi dari media sosial *Twitter* tersebut, oleh karena itu dibuatlah visualisasi untuk menganalisis kualitas penyebaran informasi dengan media sosial *Twitter*. Penelitian ini dilakukan dengan tiga tahap, yaitu *retrieving*, *preprocessing* kemudian visualisasi. Proses *retrieving* digunakan untuk mengambil data *Tweet* akun BMKG di *Twitter*. Tahap *preprocessing* berfungsi untuk mendapatkan hasil analisis kualitas informasi gempa bumi. Setelah itu hasil analisis data ketepatan waktu, relevansi, kelengkapan data beserta keakuratan data dibuat kesimpulan yang menyatakan kualitas informasi penyebaran gempa bumi, lalu hasilnya di tampilkan dalam bentuk visualisasi grafik.

Copyright © 2019 Journal of Applied Informatics and Computing.  
All rights reserved

### I. PENDAHULUAN

Indonesia merupakan negara yang memiliki letak geografis strategis. Indonesia terdiri dari beberapa pulau besar dan ribuan pulau-pulau kecil, dimana masih terdapat rangkaian gunung api yang masih aktif. Hal tersebut dapat memicu terjadinya banyak bencana alam yang tidak hanya berasal dari gunung api yang masih aktif, bisa juga berasal dari pergeseran lempeng bumi yang terjadi secara alami. Keadaan alam yang selalu berubah juga dapat memicu terjadinya banyak bencana alam yang merupakan akibat dari berbagai faktor. Dikarenakan banyaknya bencana alam yang

yang akan terjadi tanpa menunda dan bersamaan melalui beberapa saluran [2].

Seiring berkembangnya zaman, penyebaran informasi melalui teknologi saat ini sudah semakin mudah dan cepat. Salah satu cara untuk melakukan penyebaran informasi yang cepat yaitu dengan melalui media sosial. Media sosial yang masih cukup banyak pengguna khususnya di Indonesia adalah *Twitter*. Media sosial *Twitter* dapat membantu pihak BMKG dalam menyebarkan informasi peringatan dini bencana alam karena banyaknya jumlah pengguna *Twitter* di Indonesia.

*Text Mining* pada *Twitter* adalah sebuah proses yang dapat menganalisis data berupa teks dari tingkat keakuratan serta

sering mengalami gempa bumi dan letusan gunung berapi yang mengelilingi cekungan Samudra Pasifik [1].

Di Indonesia terdapat sebuah lembaga yang melaksanakan tugas pemerintahan dan bertanggung jawab dibidang meteorologi, klimatologi dan geofisika (BMKG). Peringatan dini bencana alam merupakan peringatan untuk menginformasikan kepada masyarakat tentang bencana alam

g muncul dari data tersebut secara otomatis. Dalam hal ini data tersebut adalah *Tweet* akun BMKG di *Twitter*. Dengan data yang telah diolah tersebut, hasil yang didapatkan adalah pengelompokan data untuk menganalisis kualitas informasi dari *Tweet* informasi gempa bumi BMKG di media sosial *Twitter*. Penelitian ini menggunakan pendekatan *Text Mining* dengan bantuan *Natural Language Processing* atau yang biasa dikenal dengan NLP. NLP digunakan untuk menyederhanakan data

*Tweet* agar mudah di analisis sebagai data yang relevan dan tidak relevan dalam proses *text mining*.

Selain menggunakan NLP, penelitian ini juga menggunakan algoritma pencocokan string *brute force* yang berfungsi memecahkan suatu masalah yang biasanya di dasarkan pada pernyataan masalah definisi konsep yang dilibatkan. Algoritma *brute force* memecahkan masalah dengan sangat sederhana, langsung dan dengan cara yang jelas. Di dalam penelitian ini algoritma *brute force* berperan dalam menganalisis data teks *Tweet* tingkat relevansinya.

Penelitian ini menggunakan 2 jenis visualisasi di dalam tahap penyajian data, yaitu visualisasi grafik dan peta persebaran [3]. Visualisasi grafik menampilkan kualitas penyebaran informasi gempa bumi yang dilakukan pihak BMKG di media sosial *Twitter*. Sedangkan visualisasi peta persebaran menunjukkan lokasi pengguna *Twitter* yang melakukan *retweet* terhadap *tweet* BMKG mengenai gempa bumi terjadi selama 6 bulan, dari bulan Januari hingga bulan Juni di tahun 2018.

Visualisasi di dalam penelitian ini, menunjukkan kualitas informasi atau tingkat efektivitas penyebaran informasi peringatan dini bencana alam melalui akun BMKG di media sosial *Twitter*. Data di dapatkan dari tahap pengumpulan data *tweet* akun BMKG di media sosial *Twitter*, yang telah diolah dan dianalisis. Dengan visualisasi grafik dan peta persebaran, hasil yang ditampilkan menjadi lebih jelas dan mudah di pahami.

## II. TINJAUAN PUSTAKA

### A. Twitter

*Twitter* adalah sebuah situs web yang dimiliki dan dioperasikan oleh *Twitter Inc.*, yang menawarkan jaringan sosial berupa mikroblog sehingga memungkinkan penggunaanya untuk mengirim dan membaca pesan *tweet*. *Twitter* merupakan tempat akses utama data yang menjadi topik dari penelitian ini. Dan data dari media sosial *Twitter* yang dikumpulkan dan diolah adalah:

#### 1. Tweet

*Tweet* adalah sesuatu yang ditulis user di media sosial *Twitter*. Dalam penelitian ini *Tweet* yang dikumpulkan untuk di analisis berasal dari akun BMKG yang merupakan pihak yang bertanggung jawab dalam informasi gempa bumi di Indonesia.

#### 2. Retweet

*Retweet* merupakan pengulangan dari *Tweet* user yang kemudian di tulis ulang dan bisa ditambahkan sisipan kata. *Retweet* disini menunjukkan jumlah responden dari *Tweet* BMKG.

### B. Python

*Python* merupakan salah satu bahasa pemrograman interpretative, intraktif, berbasis objek dan bisa dijalankan di semua platform [4]

Di dalam penelitian ini, *Python* digunakan sebagai bahasa pemrograman untuk membuat skrip yang berfungsi sebagai *retrieval data Tweet* akun BMKG di *Twitter*, *Text Mining data Tweet* relevan dan tidak relevan dan visualisasi. Berikut adalah *library-library* yang digunakan di dalam penelitian ini:

#### 1. Tweepy

*Tweepy* adalah *library Python* yang bertugas untuk mengakses API milik *Twitter*. *Tweepy* merupakan *library Python* yang dapat mengakses API milik *Twitter* sehingga dapat membuat *bot retrieving* data dari *Twitter* dengan skrip *Python*. *Tweepy* digunakan untuk menjembatani bahasa pemrograman *Python* dengan *Twitter* di penelitian ini. Dengan menggunakan *library Tweepy*, data *Tweet* dari akun BMKG di *Twitter* dapat di kumpulkan dan di akses untuk bahan penelitian.

#### 2. PyMySQL

*Library PyMySQL* digunakan untuk menghubungkan basis data server dengan bahasa pemrograman *Python*. Dengan adanya *library PyMySQL*, dapat memudahkan untuk melakukan pengolahan data yang ada di basis data *MySQL*.

#### 3. Unidecode

*Library Unidecode* bekerja sama pada proses *Text Mining* di dalam bahasa pemrograman *Python*. *Library Unidecode* berfungsi pada tahap *Case Folding* di dalam *Text Mining*, yaitu membuat semua huruf menjadi huruf kecil (*lower case*).

#### 4. NLTK (Natural Language Tool Kit)

*Library NTLK* di dalam penelitian ini digunakan pada tahap *Tokenizing*, yaitu berfungsi untuk memisahkan kata-kata pada setiap kalimat yang akan diolah.

#### 5. Sastrawi

*Library Sastrawi* merupakan *library* yang mempunyai peran penting di dalam proses *Text Mining*, terutama di dalam pengolahan teks Bahasa Indonesia. *Library Sastrawi* di dalam penelitian ini digunakan pada tahap *Stopword* dan *Stemming*, dimana tahap-tahap tersebut berfungsi untuk menghapus kata-kata yang tidak berhubungan dan menyederhanakan kata menjadi kata dasar.

### C. Text Mining

*Text mining* adalah penambangan yang dilakukan oleh komputer untuk menemukan informasi dan pengetahuan yang berguna yang tidak diketahui sebelumnya sehingga menemukan sesuatu yang baru yang berasal dari sumber yang berbeda-beda [5]. *Text mining* merupakan teknik yang digunakan untuk menangani masalah *classification*, *clustering*, *information extraction* dan *information retrieval*. *Text Mining* di dalam penelitian ini berfungsi untuk mengolah dokumen teks yang telah dikumpulkan di dalam basis data

yang berupa *Tweet* untuk dapat diolah ke tahapan selanjutnya. *Text Mining* memiliki 2 tahapan, yaitu:

### 1. Text Pre-processing

*Text Preprocessing* merupakan tahapan awal terhadap teks untuk mempersiapkan teks menjadi data yang dapat diolah lebih lanjut. Berikut penjelasan proses tahapan *Text Preprocessing*:

- *Case Folding*

*Case Folding* berfungsi untuk merubah semua huruf pada dokumen teks menjadi huruf kecil dan menghapus delimiter yang merupakan karakter selain huruf. Ilustrasi *case folding* yaitu mengubah semua karakter menjadi huruf kecil di tampilan pada gambar 1.

Ibukota Negara Indonesia merupakan Jakarta, sedangkan Bali adalah kota di Indonesia yang terpopuler di Dunia.	ibukota negara indonesia merupakan jakarta sedangkan bali adalah kota di indonesia yang terpopuler di dunia
--	---

Gambar 1. Ilustrasi Case Folding

- *Tokenizing*

*Tokenizing* adalah tahap pemotongan string berdasarkan pada tiap kata yang menyusunnya. Menguraikan kalimat menjadi kata per-kata di ilustrasikan pada gambar 2.

ibukota negara indonesia merupakan jakarta sedangkan bali adalah kota di indonesia yang terpopuler di dunia	"ibukota","negara" ,"indonesia", "merupakan","jaka rta", "sedangkan", "bali", "adalah", "kota", "di", "indonesia","yang" ,"terpopuler", "di", "dunia"
---	---

Gambar 2. Ilustrasi

### 2. Text Transformation

*Text Transformation* merupakan tahapan yang dipergunakan untuk menyederhanakan kata menjadi kata dasar, serta menghilangkan kata-kata yang tidak berhubungan seperti kata sambung. Dalam hal ini diperlukan *library* yang memiliki bahasa yang sesuai dengan bahasa dokumen teks yang akan diolah. Berikut tahapan-tahapan *Text Transformation*:

- *Stopword Removal / Filtering*

Tahap *Stopword Removal* berfungsi untuk menghilangkan kata-kata yang tidak penting atau kata-kata yang tidak berhubungan. Ilustrasi dari tahapan *stopword* yaitu menghilangkan kata-kata tidak penting ditampilkan pada gambar 3.

"ibukota","negara", "indonesia", "merupakan","jakarta", "sedangkan","bali", "adalah","kota", "di","indonesia","yang", "terpopuler","di", "dunia"	"ibukota","negara", "indonesia", "jakarta","bali", "kota", "indonesia", "terpopuler", "dunia"
---	---

Gambar 3. Ilustrasi Stopword Removal

- *Stemming*

*Stemming* merupakan tahap untuk menyederhanakan kata. Tahap *Stemming* berfungsi untuk mencari kata dasar pada dokumen teks yang akan diolah. Ilustrasi *stemming* yaitu menyederhanakan kata menjadi kata dasar ditampilkan pada gambar 4.

"ibukota","negara", "indonesia", "jakarta","bali", "kota","indonesia", "terpopuler", "dunia"	"ibukota","negara", "indonesia", "jakarta","bali", "kota","indonesia", "terpopuler", "dunia"
---	---

Gambar 4. Ilustrasi Stemming

### D. Algoritma Pencocokan String (*brute force*)

Algoritma *brute force* adalah sebuah pendekatan yang lempang (*straight forward*) untuk memecahkan suatu masalah, biasanya didasarkan pada pernyataan masalah (*problem statement*) definisi konsep yang dilibatkan. Algoritma *brute force* memecahkan masalah dengan sangat sederhana, langsung dan dengan cara yang jelas [6]

Cara kerja Algoritma *brute force*:

1. Mula-mula string dicocokkan pada awal teks.
2. Dengan bergerak dari kiri ke kanan, dibandingkan setiap karakter di dalam string dengan karakter yang bersesuaian di dalam teks, jika sesuai dibandingkan tersebut mengeluarkan hasil.
3. Jika string belum ditemukan kecocokan dari teks belum habis, maka geser string satu karakter ke kananda dan berulangi ke langkah 2.

Contoh dari Algoritma *brute force*:

Teks : nobody noticed him  
 Pattern : not  
 nobody **noticed** him  
 step=0 not  
 step=1 not  
 step=2 not  
 step=3 not  
 step=4 not  
 step=5 not  
 step=6 not  
 step=7 not

Pada langkah ke-7, pencarian telah berhenti karna sudah menemukan kata yang cocok dengan *pattern*. Algoritma *brute force* di dalam penelitian ini berfungsi untuk mengklasifikasikan data *Tweet* yang relevan dan yang tidak relevan untuk menganalisis tingkat efektivitas dari penyebaran informasi bencana alam gempa bumi pada akun BMKG di media sosial *Twitter*.

#### E. Kualitas Informasi

Nilai informasi ditentukan oleh banyak hal, diantaranya adalah kualitas informasi karena secara tidak langsung nilai informasi akan diperoleh. Kualitas informasi adalah “tingkat dimana informasi memiliki karakteristik isi, bentuk, dan waktu yang memberikannya nilai buat para pemakai akhir tertentu [7]. McLeod mengemukakan bahwa suatu informasi yang berkualitas harus memiliki ciri-ciri sebagai berikut [8]:

##### 1) Akurat

Informasi harus bebas dari kesalahan-kesalahan dan tidak menyesatkan. Akurat juga berarti informasi harus jelas mencerminkan maksudnya. Informasi harus akurat karena dari sumber informasi sampai ke penerima informasi kemungkinan banyak terjadi gangguan (*noise*) yang dapat mengubah informasi tersebut.

Di dalam penelitian ini, untuk tingkat akurasi dari data *Tweet* sudah terpenuhi, dikarenakan data di ambil langsung dari *Tweet* pihak BMKG di media sosial *Twitter* dan merupakan badan resmi penyebaran informasi bencana alam di Indonesia yang memiliki metode sendiri untuk meneliti bencana alam khususnya gempa bumi yang terjadi di Indonesia.

##### 2) Tepat Waktu

Informasi yang datang pada si penerima tidak boleh terlambat. Informasi yang sudah usang tidak akan mempunyai nilai lagi karna informasi merupakan landasan dalam pengambilan keputusan. Informasi tepat waktu harus tersedia atau ada pada saat informasi tersebut diperlukan, tidak besok atau beberapa jam lagi, sehingga bisa disimpulkan bahwa informasi bencana alam

khususnya gempa bumi yang tepat waktu tidak lebih dari satu jam. Di dalam penelitian ini, untuk menentukan *Tweet* informasi gempa bumi yang tepat waktu yaitu waktu terjadinya gempa yang sudah ditambahkan satu jam lalu dibandingkan dengan waktu ketika *Tweet* dibuat. Apabila waktu *Tweet* lebih kecil dibandingkan waktu terjadinya gempa yang sudah ditambahkan satu jam, maka *Tweet* tersebut merupakan informasi yang tepat waktu.

##### 3) Relevan

Relevan artinya, Informasi yang diberikan harus sesuai dengan yang dibutuhkan. Relevansi informasi untuk orang satu dengan orang yang lain berbeda. Relevan di dalam penelitian ini adalah, apabila informasi *Tweet* mengandung kata gempa bumi di dalamnya.

##### 4) Lengkap

Lengkap ialah tidak boleh ada bagian informasi yang penting atau esensial bagi pengambil keputusan atau pelaksanaan tugas yang hilang, karena menghasilkan keputusan yang salah nantinya. Di dalam penelitian ini, informasi *Tweet* yang lengkap mengandung informasi gempa bumi yang berisikan data-data penting di dalamnya.

Dalam penelitian ini, menggunakan tabel standar efektivitas untuk menentukan tingkat kualitas informasi khususnya efektivitas dari hasil pengolahan data penyebaran informasi gempa bumi akun BMKG di *Twitter* dalam bentuk persentase.

TABEL I.  
TABEL STANDAR EFEKTIFITAS

Rasio Efektivitas	Tingkat Capaian
< 40%	Sangat Tidak Efektif
40% - 59.99%	Tidak Efektif
60% - 79.99%	Cukup Efektif
> 80%	Sangat Efektif

#### F. Visualisasi

Visualisasi adalah rekayasa dalam pembuatan gambar, diagram atau animasi bahkan pemetaan untuk penampilan suatu informasi.

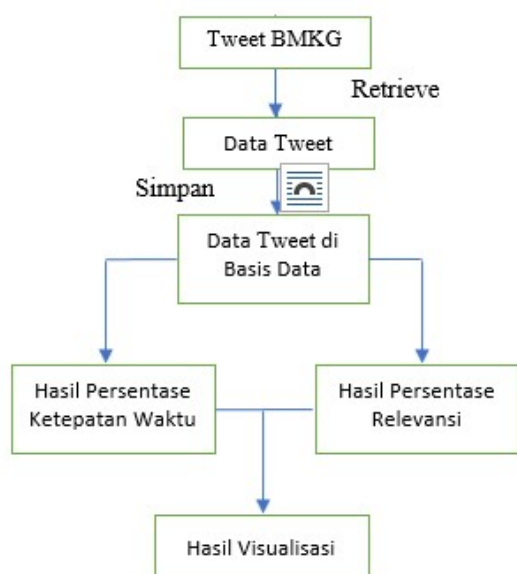
- 1) Metode penggunaan komputer untuk mentransformasi simbol menjadi geometrik.
- 2) Memungkinkan peneliti mengamati simulasi dan komputasi.
- 3) Memberikan cara untuk melihat yang tidak terlihat
- 4) Memperkaya proses penemuan ilmiah dan mengembangkan pemahaman yang lebih dalam dan tak diduga

- 5) Dalam berbagai bidang telah merevolusikan cara ilmuwan meneliti sains.

Penelitian ini menggunakan 2 jenis visualisasi, yaitu visualisasi grafik dan peta persebaran.

### III. METODE PENELITIAN

Penelitian yang dilakukan ini dibagi menjadi 3 tahap, tahapan yang pertama yaitu tahap pengumpulan data. Setelah data dikumpulkan tahapan selanjutnya yaitu pengolahan data dimana data diolah menjadi data yang lebih berkualitas. Setelah itu tahap penyajian data yang berisikan visualisasi dari data olahan.



Gambar 5. Tahapan Proses Penelitian

Gambar 5 menunjukkan tahapan proses di dalam penelitian ini secara keseluruhan. Dimana pada tahapan awal yang dilakukan adalah melakukan *retrieve* dari *Tweet* akun BMKG di *Twitter* dan langsung disimpan ke Basis Data. Data *Tweet* yang ada di Basis Data diolah dan dianalisis hingga hasil yang di dapatkan adalah persentase ketepatan waktu. Dan data *Tweet* yang ada di Basis Data dengan melakukan pengolahan data teks dan analisis sehingga hasil yang di dapatkan adalah persentase relevansi data *Tweet*. Dari hasil yang telah di dapatkan dibuat visualisasi agar hasilnya lebih mudah dimengerti.

#### A. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini yaitu observasi. Langkah-langkah yang dilakukan dalam observasi sebagai berikut:

1. Melakukan *retrieving* data *Tweet* dari akun BMKG di *Twitter* dengan menggunakan skrip *Python* dan bantuan

*Library Tweepy*.

Variabel-variabel yang dipelajari dalam *retrieving* data *Tweet* yaitu:

- **ID**  
Nomor id dari *Tweet* BMKG di *Twitter*
- **Text Tweet**  
Isi dari teks *Tweet* akun BMKG di media sosial *Twitter*. Digunakan untuk mengolah dan mendapatkan hasil analisis data.
- **Retweet\_count**  
Jumlah *retweet* dari *Tweet* BMKG di *Twitter*
- **Favourite\_count**  
Jumlah *favourite* dari *Tweet* BMKG di *Twitter*
- **Date\_time**  
Merupakan tanggal dari *Tweet* dibuat. Digunakan untuk mengetahui tanggal dan waktu ketika di publikasi.

2. Menyimpan data-data *Tweet* yang telah di dapatkan lalu disimpan ke dalam basis data.

no	text	retweet_count	favourite_count	date_time
576	b#Gempa Mag 5.1, 14-Jun-18 06:06:10 WIB, Lok:1.90...	73	61	2018-06-13 23:12:20
577	b#Gempa Mag 5.1 SR, 14-Jun-18 06:06:10 WIB, Lok:1...	34	38	2018-06-13 23:11:10
578	b#Peringatan dini cuaca wilayah Banten [14 Juni 20...	4	18	2018-06-13 21:00:24
579	b#Peringatan dini cuaca wilayah Banten [14 Juni 20...	7	18	2018-06-13 18:25:48
580	b#Peringatan dini cuaca wilayah JABODETABEK [14 Ju...	14	18	2018-06-13 17:28:57
581	b#Siaran Pers #BMKG: Gempabumi M=4.8 Mengguncang S...	93	72	2018-06-13 15:48:05
582	b#Siaran Pers #BMKG: Update 9 Kali Gempabumi Susul...	158	100	2018-06-13 15:08:00
583	b#Peringatan dini cuaca wilayah Kalimantan Utara [...]	0	8	2018-06-13 14:25:43
584	b#Gempa Mag 4.8, 13/06/2018 20:06:40 (Pusat gempa...	64	66	2018-06-13 13:41:07
585	b#Gempa Mag 4.8, 13-Jun-18 20:06:40 WIB, Lok:6.88...	113	75	2018-06-13 13:39:13

Gambar 6. Data *Tweet* di Basis Data

Gambar 6 menunjukkan contoh data yang di dapatkan dalam proses pengambilan data *Tweet* BMKG di *Twitter*. No pada gambar merupakan nomor urutan pengumpulan data, *text* merupakan isi dari *Tweet*, *retweet\_count* merupakan jumlah *Retweet*, *favourite\_count* merupakan jumlah *Favourite* dan *date\_time* merupakan waktu dari tweet tersebut di buat

#### B. Teknik Pengolahan dan Analisis Data

Tahap-tahap yang dilakukan untuk pengolahan data dan analisis data adalah sebagai berikut:

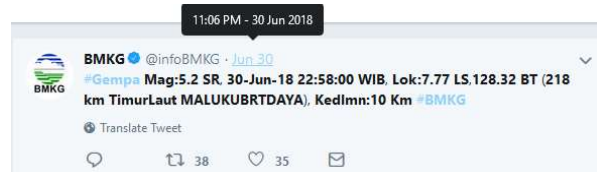
1. Mengolah data *Tweet* yang tersimpan di dalam basis data dengan skrip pengolahan data dan *library* penghubung.
2. Melakukan analisis pengolahan data untuk mengetahui ketepatan waktu dari informasi *Tweet* gempa bumi. Pengolahan data yang dilakukan yaitu membandingkan waktu terjadinya gempa dengan waktu *Tweet* dibuat. Dikarenakan kriteria informasi yang tepat waktu tidak lebih dari beberapa jam, sehingga apabila waktu *Tweet* dibuat, tidak lebih satu jam dari waktu terjadinya gempa, maka *Tweet* tersebut termasuk informasi yang tepat waktu.



Gambar 5. Tweet Informasi Gempa Tepat Waktu

Gambar 7 menunjukkan contoh *Tweet* informasi gempa bumi yang dilakukan secara tepat waktu, sehingga *user* yang mendapatkan informasi bisa dengan segera melakukan tindakan mitigasi bencana gempa bumi.

3. *Text Mining* digunakan untuk menyederhanakan teks *Tweet* agar mudah dianalisis menggunakan *Natural Language Processing*. Proses *text mining* secara umum terbagi 2, yaitu:
  - *Text Processing*  
Tindakan yang dilakukan pada tahap ini adalah *toLowerCase*, yaitu mengubah semua karakter huruf menjadi huruf kecil dan *Tokenizing* yang berfungsi sebagai penguraian deskripsi yang semula berupa kalimat menjadi kata-kata lalu menghilangkan delimiter-delimiter seperti tanda titik (.), koma (,), spasi dan karakter angka.
  - *Featured Selection*  
Tindakan yang dilakukan adalah *stopword removal* yaitu menghilangkan kata-kata yang tidak berhubungan seperti kata sambung dan *stemming* yaitu melakukan penyederhanaan kata menjadi kata dasar terhadap isi teks *Tweet* yang dianalisis.
4. Menggunakan algoritma pencocokan *string* (*brute force*) untuk menganalisis kelengkapan dan relevansi dari data teks *Tweet*. Setelah data teks *Tweet* melewati tahapan *text mining*, lalu dilakukan pencocokan dengan kriteria data teks yang lengkap dan relevan untuk menganalisis tingkat kelengkapan dan relevansi dari isi teks *Tweet*.



Gambar 6. Tweet Informasi Gempa yang Baik

Gambar 8 menunjukkan contoh *Tweet* relevan terhadap *Tweet* informasi gempa bumi dan mengandung informasi yang lengkap. Dapat dilihat pada gambar, *Tweet* masih mengandung informasi penting mengenai informasi gempa bumi, seperti:

- Menampilkan kata “gempa” di dalam isi teks *Tweet* yang menunjukkan informasi tersebut termasuk dalam kategori bencana alam gempa bumi. Data ini digunakan untuk menentukan relevansi *Tweet*

informasi tersebut dengan informasi gempa bumi.

- Menampilkan informasi kekuatan gempa, yaitu “Mag:5.2”
  - Menampilkan waktu kejadian dari gempa yang akan terjadi
  - Menampilkan lokasi titik gempa dalam bentuk satuan lintang dan bujur
  - Menampilkan informasi daerah terjadinya gempa
  - Menampilkan kedalaman gempa
5. Visualisasi data yang telah dikumpulkan dari hasil analisis ke dalam bentuk grafik dan peta persebaran.
  6. Hasil visualisasi grafik menampilkan hasil analisis relevansi, ketepatan waktu dan kelengkapan data dalam kurun waktu 6 bulan, dari bulan Januari hingga Juni tahun 2018.
  7. Hasil visualisasi peta persebaran yang di tampilkan merupakan peta yang dibuat oleh pihak BMKG. Peta menampilkan titik-titik gempa yang terjadi dalam kurun waktu 6 bulan, dari bulan Januari hingga bulan Juni tahun 2018
  8. Dari visualisasi grafik dan peta persebaran dapat disimpulkan kualitas penyebaran informasi dari *Tweet* akun BMKG di *Twitter*.

### C. Teknik Penyajian Data

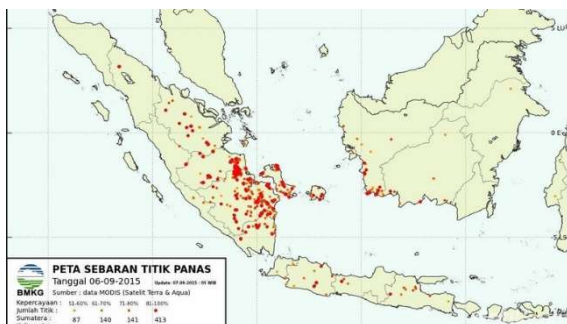
Penyajian data yang digunakan yaitu melakukan visualisasi data ke dalam bentuk grafik dan peta persebaran agar lebih mudah dipahami.



Gambar 9. Grafik Pie (Direktorat Jenderal Perikanan Tangkap, 2013)

Gambar 9 merupakan contoh penyajian data dari hasil visualisasi dalam bentuk grafik *pie*. Digunakan untuk memvisualisasikan kualitas informasi *Tweet* mengenai gempa bumi yang dilakukan oleh akun BMKG di media sosial *Twitter*.





Gambar10. Peta Persebaran (BMKG, 2015)

Gambar 10 merupakan contoh penyajian data visualisasi dalam bentuk peta persebaran. Digunakan untuk menampilkan titik informasi *Tweet* mengenai gempa bumi yang dilakukan oleh akun BMKG di media sosial *Twitter*.

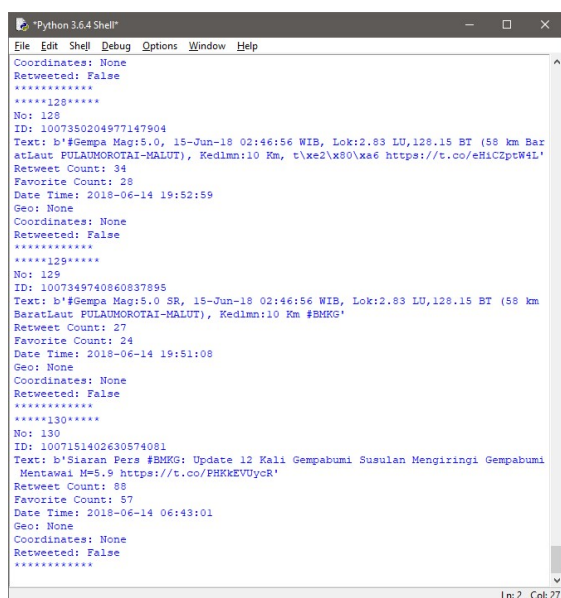
#### IV HASIL DAN PEMBAHASAN

### A. Hasil Implementasi

Hasil Implementasi dibagi menjadi 3 tahap yaitu retrieving data, preprocessing data dan visualisasi.

### 1) Retrieving Data

Pada tahap ini dilakukan *retrieving* data *Tweet* dan *Rewet* dari akun BMKG (@infoBMKG) di media sosial *Twitter* untuk mendapatkan data *Tweet* BMKG. Tahap *Retrieving* dilakukan dengan menjalankan skrip *Retrieving Timeline.py*.



Gambar 11. Hasil Retrieving Data Tweet

Gambar 11 menunjukkan hasil data dari proses *retrieving* data dari akun BMKG (@infoBMKG) di media sosial *Twitter* dan disimpan ke Basis Data.

no	id	text	retweet_count	favourite_count	date_time
1	1013214786916503562	b#Gempa Mag 3.7, 01-Jul-18 06:18:35 WIB, Lok 8.14...	45	29	2018-07-01 00:16:44
2	1013214647397126145	b#Gempa Mag 3.7, 01/07/2018 06:18:35 (Pusat gempa...	13	15	2018-07-01 00:16:11
3	1013091325481283584	b#Gempa Mag 5.2 SR, 30- Jun-18 22:58:00 WIB, Lok 7.77...	36	32	2018-06-30 16:06:08
4	1013090900950507522	b#Gempa Mag 5.2, 30- Jun-18 22:58:00 WIB, Lok 7.77...	68	51	2018-06-30 16:04:27
5	1012868605275451393	b#Gempa Mag 3.3, 30/06/2018 03:03:58 (Pusat gempa...	18	18	2018-06-30 01:21:08
6	1012856082048286720	b#Gempa Mag 3.3, 30- Jun-18 03:03:58 WIB, Lok 1.94...	22	20	2018-06-30 00:31:22
7	1012814118867226624	b#Gempa Mag 3.3, 30- Jun-18 03:03:58 WIB, Lok 1.94...	39	27	2018-06-29 21:44:37
8	1012801927615148032	b#Gempa Mag 3.3, 30/06/2018 03:03:58 (Pusat gempa...	16	25	2018-06-29 20:56:11

Gambar 12. Data *Retrieving Tweet* di Basis Data

Gambar 12 menunjukkan hasil data *retrieving* dari akun BMKG (@infoBMKG) di media sosial *Twitter* yang telah disimpan ke Basis Data.

## 2) Prepossessing Data

Pada tahap ini dilakukan *Prepossesing* data dari data yang telah dikumpulkan di basis data untuk mengolah dan mengimplementasikan metode *Text Mining*. Pada tahap ini, setelah data diolah dan disederhanakan agar mudah di pahami menggunakan NLP, data dicocokkan dengan *pattern* menggunakan algoritma pencocokkan string (*brute force*).

*Pattern-pattern* yang telah disediakan sebelumnya mengandung beberapa poin penting di dalam *Tweet* informasi gempa bumi, yaitu:

- Terdapat kata “gempa” di dalamnya. Menunjukkan bahwa informasi tersebut merupakan informasi mengenai gempa bumi. Data ini digunakan untuk menganalisis relevansi data *Tweet* dengan informasi gempa bumi.
- Besar kekuatan gempa.
- Waktu terjadinya gempa yang berfungsi sebagai memperingatkan *user* untuk dapat bersiap-siap menghadapi gempa bumi. Data ini diolah untuk dibandingkan dengan waktu *Tweet* sehingga didapatkan persentase ketepatan waktu.
- Titik koordinat terjadinya gempa
- Daerah di Indonesia yang tertimpa musibah gempa
- Kedalaman gempa yang terjadi

Poin-poin yang berisikan informasi penting gempa ini digunakan sebagai pattern pembandingan untuk menentukan tingkat relevansi data dan kelengkapan informasi.

### 3) Visualisasi

Pada tahap ini dilakukan proses visualisasi dari data hasil *Prepossesing* ke dalam bentuk grafik. Implementasi perancangan input dan output data yang diperlukan yaitu:

- Menjalankan skrip “Retrieve\_Timeline.py” untuk melakukan *retrieving* data dari akun BMKG di media sosial *Twitter*.
- Menjalankan skrip “Relevance & Complete.py” untuk melakukan proses *Prepossesing* yang melakukan analisis

terhadap data yang telah disimpan di Basis Data untuk mendapatkan hasil data tingkat relevansi dan kelengkapan data.



Gambar 13. Hasil Preprocessing Data Relevan

Gambar 13 menunjukkan hasil dari *Preprocessing* data yang relevan. Ketika teks yang sudah disederhanakan, dilakukan pengolahan data menggunakan algoritma pencocokan string (*brute force*) pada *pattern* yang sudah disediakan, teks *Tweet* tersebut mengandung poin informasi penting dari *Tweet* informasi gempa bumi, yaitu:

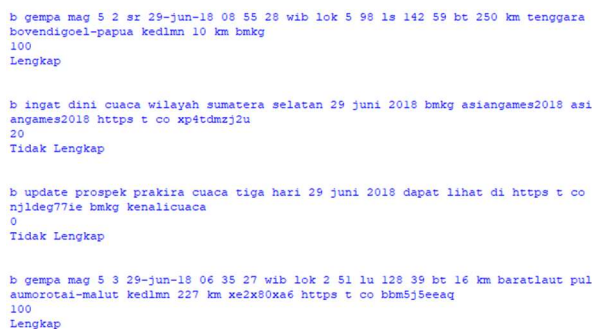
- 1) Terdapat kata “gempa”.

Kata gempa menentukan relevan atau tidaknya data tersebut.



Gambar 14. Hasil Perhitungan Data Relevan

Gambar 14 menampilkan hasil dari perhitungan data yang relevan sebesar 927 dari 927 data, dan persentase yang didapatkan sebesar 100%. Hasil persentase 100% didapatkan karena data yang diolah hanya data *Tweet* informasi gempa bumi BMKG di *Twitter*. Sedangkan untuk nilai kualitas informasi, dengan jumlah persentase data relevan dari kualitas informasi sebesar 25%. Dan dari hasil persentase data relevan 100%, maka di dapatkan hasil persentase sebesar 25%.

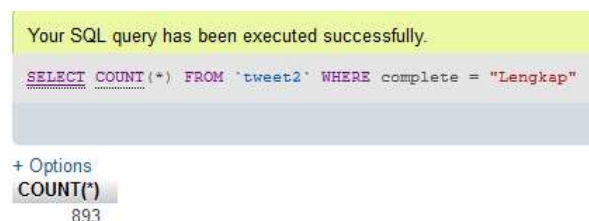


Gambar 15. Hasil Preprocessing Data Lengkap

Gambar 15 menunjukkan hasil dari *Preprocessing* data yang relevan. Ketika teks yang sudah disederhanakan, dilakukan pengolahan data menggunakan algoritma pencocokan string (*brute force*) pada *pattern* yang sudah disediakan, teks *Tweet* tersebut mengandung beberapa poin informasi penting dari *Tweet* informasi gempa bumi, yaitu:

- 2) Terdapat informasi kekuatan gempa (*magnitude*)
- 3) Terdapat informasi waktu terjadinya gempa bumi
- 4) Terdapat informasi titik koordinat gempa (*epicenter*) dengan satuan Lintang dan Bujur.
- 5) Terdapat informasi daerah terjadinya gempa di Indonesia
- 6) Terdapat informasi kedalaman gempa

Minimal terdapat 3 poin dari 5 poin penting untuk menentukan data tersebut merupakan data Lengkap. Apabila pada teks yang diolah mengandung 5 poin penting informasi gempa bumi, hasil yang didapatkan adalah data tersebut Lengkap dan apabila kurang dari 3 poin, hasilnya adalah data tersebut Tidak Lengkap.

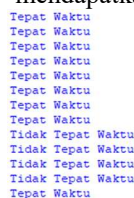


Gambar 16. Hasil Perhitungan Data Lengkap

Gambar 16 menampilkan hasil perhitungan data yang lengkap sebesar 893 dari 927 data, dan persentase yang didapatkan sebesar 96.33%.

Sedangkan untuk nilai kualitas informasi, dengan jumlah persentase data lengkap dari kualitas informasi sebesar 25%. Dan dari hasil persentase data relevan 96.33%, maka di dapatkan hasil persentase sebesar 24.08 %.

- 1) Menjalankan skrip “On\_Time.py” untuk melakukan proses *Preprocessing* yang melakukan analisis terhadap data yang telah disimpan di Basis Data untuk mendapatkan hasil data ketepatan waktu.



Gambar 17. Hasil Preprocessing Data Tepat Waktu

Gambar 17 menunjukkan hasil dari *Preprocessing* data ketepatan waktu. Dengan melakukan perbandingan data yaitu waktu *Tweet* dibuat dan waktu terjadinya gempa dengan ditambahkan satu jam.





Gambar 18. Hasil Perhitungan Data Tepat Waktu

Gambar 18 menampilkan hasil dari perhitungan data yang tepat waktu sebesar 742 dari 927 data, dan persentase yang didapatkan sebesar 80.04%.

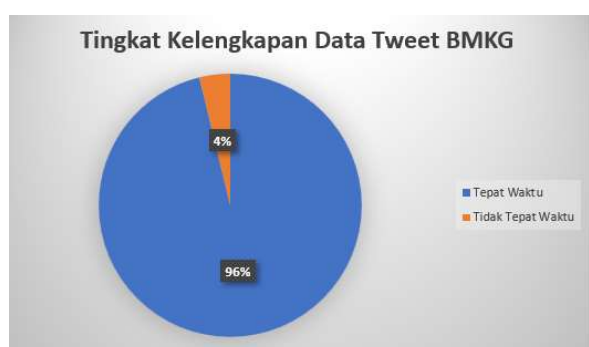
Sedangkan untuk nilai kualitas informasi, dengan jumlah persentase data tepat waktu dari kualitas informasi sebesar 25%. Dan dari hasil persentase data tepat waktu sebesar 80.04%, maka di dapatkan hasil persentase sebesar 20.01%.

### B. Pembahasan

Berdasarkan implementasi dan pengujian yang di lakukan menghasilkan beberapa poin penting terhadap kualitas penyebaran informasi akun BMKG di media sosial *Twitter*, antara lain:



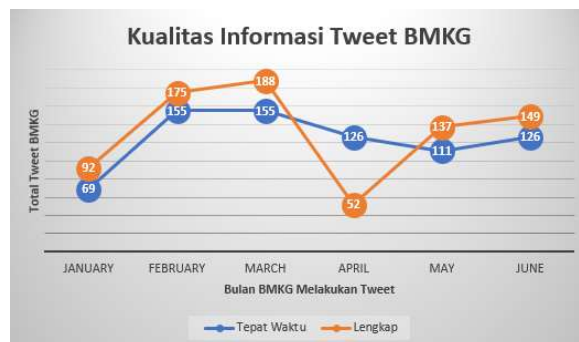
Gambar 19 Grafik Pie Ketepatan Waktu



Gambar 20. Grafik Pie Kelengkapan Data

Gambar 19 menunjukkan hasil pengolahan data untuk tingkat ketepatan waktu *Tweet* informasi BMKG. Setelah data diolah didapatkan hasil persentase sebesar 80% untuk data yang tepat waktu dan 20% data yang tidak tepat waktu. Gambar 20 menunjukkan hasil pengolahan data untuk tingkat kelengkapan data *Tweet* informasi BMKG. Setelah data

diolah didapatkan hasil persentase sebesar 96% untuk data yang lengkap dan 4% data yang tidak lengkap.



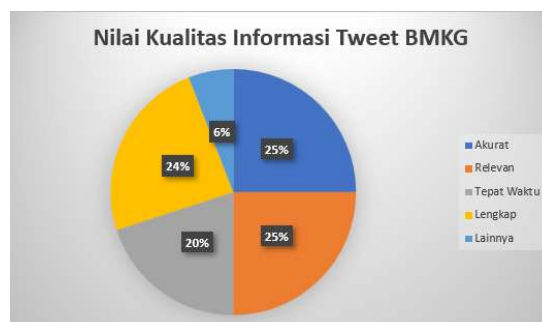
Gambar 21. Hasil Kualitas Informasi Januari-Juni Tahun 2018

Gambar 21 menunjukkan hasil Kualitas Informasi dari *Tweet* BMKG di media sosial *Twitter*. Hasil menampilkan tingkat ketepatan waktu dan tingkat kelengkapan data yang ditampilkan dalam kurun waktu 6 bulan dari bulan Januari Hingga Juni di tahun 2018.

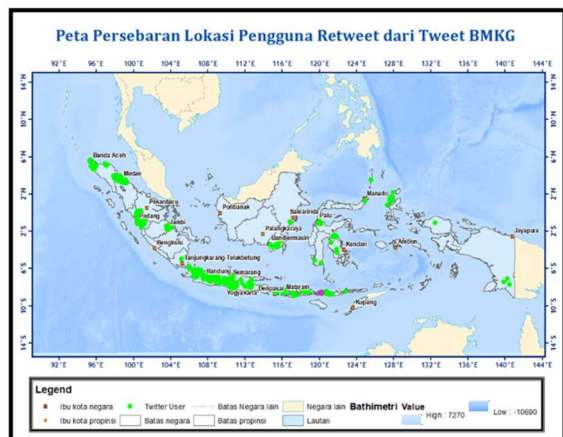
Hasil pengolahan data secara keseluruhan ditampilkan pada gambar 22. Didapatkan tingkat Kualitas Informasi dari *Tweet* informasi BMKG mengenai gempa bumi di Indonesia melalui media sosial *Twitter*, dengan hasil yang didapatkan adalah sebagai berikut:

1. Tingkat Akurasi 25%
2. Tingkat Relevansi 25%
3. Tingkat Ketepatan Waktu 20.01%
4. Tingkat Kelengkapan Data 24.08%

Sehingga data hasil Kualitas Informasi *Tweet* BMKG didapatkan sebesar 94.09% dan jika dibandingkan dengan tabel Kualitas Informasi, hasil yang didapatkan adalah Kualitas Informasi dari *Tweet* Informasi BMKG "Sangat Baik".



Gambar 22. Grafik Pie Hasil Kualitas Informasi



Gambar 23. Peta Persebaran *Retweet* Terhadap *Tweet* BMKG

Gambar 23 menunjukkan hasil visualisasi peta persebaran lokasi pengguna yang melakukan *retweet* terhadap *tweet* BMKG mengenai gempa bumi di Indonesia yang terjadi pada bulan Januari hingga Juni di tahun 2018. Simbol warna jingga menunjukkan ibu kota negara dan ibu kota propinsi yang ada di Indonesia. Simbol warna hijau menunjukkan titik lokasi pengguna yang melakukan *retweet* terhadap *tweet* BMKG. Data didapatkan dari proses pengumpulan data lokasi pengguna yang melakukan *retweet* terhadap *tweet* BMKG, setelah itu pemetaan peta persebaran dibuat berdasarkan data lokasi yang telah dikumpulkan.

## V. KESIMPULAN

Berdasarkan hasil yang di dapatkan dari pengujian dan analisis hasil visualisasi maka dapat di simpulkan bahwa:

1. Kualitas penyebaran informasi gempa bumi di media sosial *Twitter* yang telah di lakukan BMKG selama 6 bulan terakhir dari bulan Januari hingga bulan Juni di tahun 2018 sudah “Baik” dan efektif dikarenakan hasil kualitas informasi yang didapatkan sebesar 94.09%.
2. Kualitas penyebaran informasi menunjukkan bahwa:
  - BMKG melakukan *Tweet* informasi gempa bumi BMKG di *Twitter* di bulan Januari hingga Juni tahun 2018 sebanyak 927 *Tweet*.
  - *Tweet* dengan persentase ketepatan waktu informasi gempa bumi tertinggi yaitu sebanyak 155 *Tweet* terjadi di bulan Februari dan Maret.

- Tingkat dengan kelengkapan data tertinggi terjadi di bulan Maret dengan total *Tweet* sebanyak 188 *Tweet*.

3. Hasil pengolahan data dapat di visualisasikan dalam bentuk visualisasi grafik dan peta persebaran yang berisikan total *Tweet* informasi gempa dan hasil analisis kualitas informasi *Tweet* BMKG berserta lokasi pengguna yang melakukan *retweet* terhadap *tweet* BMKG mengenai gempa bumi selama 6 bulan dari bulan Januari hingga bulan Juni di tahun 2018.

## UCAPAN TERIMA KASIH

Terimakasih kepada segenap Manajemen Politeknik Negeri Batam yang telah mendukung penulis dalam melakukan Riset ini.

## DAFTAR PUSTAKA

- [1] Pusdalops (2016) Ring of Fire. [Online]. Tersedia: <http://bpbd.jabarprov.go.id/index.php/component/k2/item/38-cincin-api-pasifik> (diakses 8 Maret 2018)
- [2] A. T. Chatfield and U. Brajawidagda, “Twitter early tsunami warning system: A case study in Indonesia’s natural disaster management,” in 2013 46th Hawaii International Conference on System Sciences, 2013, pp. 2050–2060.
- [3] D. E. Kurniawan, and A. Fatulloh, “Clustering of Social Conditions in Batam, Indonesia Using K-Means Algorithm and Geographic Information System,” International Journal of Earth Sciences and Engineering (IJE) 2017, Vol.10 No.05
- [4] R. Mareta, A. D. Rahmaningsih, and R. D. Firmansyah, “Pendeteksi Ketinggian Air Interaktif Dengan Aplikasi Telegram Berbasis Raspberry Pi,” JST (Jurnal Sains dan Teknol., vol. 6, no. 2, pp. 279–289, 2017.. 2012
- [5] E. M. G. Younis, “Sentiment analysis and text mining for social media microblogs using open source tools: an empirical study,” Int. J. Comput. Appl., vol. 112, no. 5, 2015.
- [6] B. W. Santoso, F. Sundawa, and M. Azhari, “Implementasi Algoritma Brute Force Sebagai Mesin Pencari (Search Engine) Berbasis Web Pada Database,” J. Sisfotek Glob., vol. 6, no. 1, 2016.
- [7] J. O’Brien and G. Marakas, Introduction to information systems. McGraw-Hill, Inc., 2009.
- [8] R. McLeod and G. Schell, Sistem informasi manajemen. Indeks, 2004. Terjemahan oleh Ali Akbar Yulianto dan Afia R. Fitriati, Salemba Empat, Jakarta 2008